

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) EP 0 957 367 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
17.11.1999 Bulletin 1999/46

(51) Int. Cl.⁶: G01R 29/00, H04N 17/00

(21) Application number: 98400917.5

(22) Date of filing: 14.04.1998

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(72) Inventor: Le Clerc, François
35700 Rennes (FR)

(74) Representative:
Hartnack, Wolfgang, Dipl.-Ing. et al
Deutsche Thomson-Brandt GmbH
Licensing & Intellectual Property,
Karl-Wiechert-Allee 74
30625 Hannover (DE)

(71) Applicant: THOMSON multimedia
92648 Boulogne Cédex (FR)

(54) Method for estimating the noise level in a video sequence

(57) In noise measurement for video sequences it is difficult to distinguish between picture content and noise. In order to improve the measurement reliability the results (σ_{p1} , σ_{p2} , σ_{p3}) of two different noise level computing methods are combined. One computation relies on the analysis of displaced field or frame differences (DFD), the other is based on the values of the field or frame differences (FD) over static picture areas.

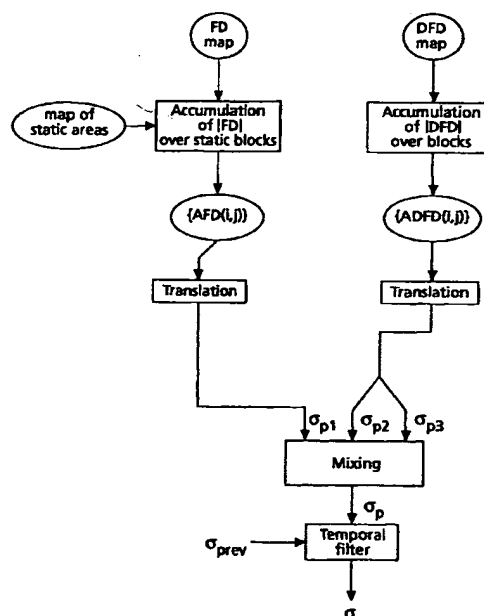


Fig.2

EP 0 957 367 A1

Description

[0001] The invention relates to a method for estimating the noise level in a video sequence.

5 Background

[0002] EP-A-0562407 discloses a method of noise measurement in conjunction with a block-matching motion estimation algorithm, the principle of which is to derive a noise level from the minimum of accumulated absolute pixel difference values, leading to a displaced field or frame differences (DFD) value, the accumulation taking place over
10 predetermined pixel blocks.

A paper by Q. Zhang and R. Ward, entitled "Automatic assessment of signal-to-thermal noise ratio of television images", Vol. 41, No. 1, IEEE Transactions on Consumer Electronics (February 1995), discloses a method for measuring the noise level from the TV pictures as such. This method is based on the application of a two-dimensional highpass filter on the images in order to remove the majority of the (non-noisy) image content. Thereafter the smoothest regions
15 of the picture, i.e. those having minimum energy with respect to brightness variations, are selected and the noise power is estimated from their remaining average power.

That paper says that in digital image processing the customary procedure to estimate the level of thermal noise in the image is to analyse "smooth regions, i.e. regions containing constant luminance (grey levels)".

20 Invention

[0003] The method described in EP-A-0562407 lacks robustness because it is based solely on the minimum of the distribution of estimates for each block of the picture, and therefore depends on the shape and deviation of this distribution. The method described by Zhang et al. suffers from the same shortcoming, as the computation of the noise level
25 is eventually based on the low-end tail of the distribution of noise energies over subimages of a picture. Thus, for pictures with few areas of high spatial frequencies, there is a risk of under-estimating the noise level.

The proposed method alleviates this problem by biasing the estimation largely on averages, rather than minima, of noise energy measurements. The measurement performed over static areas, in particular, is independent of the spatial frequency contents of the picture.

[0004] It is one object of the invention to disclose a method for more reliable noise estimation. This object is achieved by the method disclosed in claim 1.

[0005] In the invention, additional motion information provided by e.g. a motion-compensated interpolation is used in order to compute a more robust and accurate estimation of the noise level in a video sequence. Ideally, if motion estimation is error-free, the remaining differences between the grey levels of input pixels from the two source picture blocks
35 put in correspondence by an estimated motion vector must be the result of noise.

The additional motion information may also be derived from motion vector information of an MPEG bitstream.

[0006] Modifying the field or frame rate of a video sequence by interpolating pictures located temporally between the source pictures is required for picture rate upconversion or standards conversion. The best conversion quality is achieved if the motion of objects in the source sequence is estimated and used to interpolate each pixel along the direction
40 of its associated motion vector. Another application of this technique is noise reduction by means of a temporal filter, with the goal of improving either the picture quality or the coding efficiency, e.g. of MPEG2 encoders.

Motion estimation can be performed by finding the vectors that provide the best match between pixels or blocks of pixels mapped from a previous or current picture to a next picture. The mathematical criterion used for the selection of a motion vector is usually the minimisation of the sum of the absolute values of the displaced field difference or displaced
45 frame difference of a pixel block, as described in Fig. 1. An intermediate field or frame IF to be interpolated is located temporally between a previous field or frame PF and a next field or frame NF.

The temporal distance between PF and NF is T, between PF and IF $\alpha \cdot T$, and between IF and NF $(1-\alpha) \cdot T$.

The zero vector $\underline{0} = (0,0)$ passes through points $I^P(x,y)$ in PF, $I(x,y)$ in IF, and $I^N(x,y)$ in NF. A current candidate motion vector $\underline{v} = (v_x, v_y)$ passes through points $I^P(x-\alpha \cdot v_x, y-\alpha \cdot v_y)$ in PF, $I(x,y)$ in IF, and $I^N(x+(1-\alpha) \cdot v_x, y+(1-\alpha) \cdot v_y)$ in NF.

50 The frame difference (for vector $\underline{0}$) is

$$FD = I^N(x,y) - I^P(x,y).$$

The displaced frame difference for vector \underline{v} is

55

$$DED(\underline{v}) = I^N(x+(1-\alpha) \cdot v_x, y+(1-\alpha) \cdot v_y) - I^P(x-\alpha \cdot v_x, y-\alpha \cdot v_y).$$

[0007] The interpolation of the output pictures is carried out along the direction of the estimated motion vectors. The

quality of the interpolation is limited by the accuracy of the motion vectors, except in static parts of the pictures where the motion is known to be exactly zero. It is therefore advantageous to detect static areas in the source images and to implement a specific interpolation mode for moving pixels, thereby optimising the interpolation output resolution. A specific solution for detecting such static areas is disclosed in another application of the applicant, internal reference PF980013, filed at the same date.

[0008] The inventive noise level estimation, however, is based on source pictures only. Therefore, if Fig. 1 is applied to the noise level estimation, intermediate field or frame IF is that current source picture for which the noise level is to be estimated.

According to the invention, the results of two different noise level computing methods can be combined in order to improve the reliability of the noise level estimation. One computation relies on the analysis of DFDs, the other is based on the values of the field or frame differences (FD) over static areas.

[0009] The availability of an accurate estimate of the noise level potentially improves the performance of many image processing algorithms in the presence of noise because it allows to adapt the algorithm parameters and thresholds to that noise level. Applications include: motion estimation, noise reduction, detection of static areas, film mode and film phase detection, detection of cuts, and many others.

[0010] In principle, the inventive method is suited for estimating the noise level for a current source field or frame of a video sequence, based on the differences between pixel values of blocks in a previous field or frame and corresponding pixel values of corresponding blocks in a future field or frame, wherein either said previous or said future field or frame can be said current field or frame itself, and wherein at least one block of each corresponding couple of blocks is a motion-compensated pixel block or is mapped to the other block by an associated motion vector estimate.

In addition, static picture areas can be determined and the differences between pixel values of blocks in a static picture area of a previous field or frame and corresponding pixel values of corresponding blocks in a future field or frame can be used to estimate a further noise level estimate which is then combined with said noise level estimate in order to form a final noise level estimate, wherein said previous and/or said future field or frame used for the evaluation of said differences between pixel values of a block in a static picture area can be different from said previous and/or said future field or frame used for the evaluation of the differences concerning said motion-compensated pixel blocks or said mapped blocks.

[0011] Advantageous additional embodiments of the inventive method are disclosed in the respective dependent claims.

Drawings

[0012] Embodiments of the invention are described with reference to the accompanying drawings, which show in:

Fig. 1 picture to be interpolated between and from a previous source picture and a next source picture, or, current source picture, between a previous source picture and a next source picture, for which the noise level is to be estimated;

Fig. 2 flow chart for the inventive noise level computation.

Embodiments

[0013] Input data to the inventive noise level estimation (in one field or frame hereafter referred to as the current field or frame) include:

- a map of displaced field or frame differences which may be a by-product of a motion estimation;
- a map of the input pixels or blocks of pixels which have been detected as being non-moving;
- a map of field or frame differences, computed between a previous frame and a next frame, located in time respectively before and after the current frame, if the source images are progressive ones, or in case of interlaced source images computed between a previous field and a next field located in time respectively before and after the current field, with the constraint that said previous and next fields have the same parity, i.e. that both are top fields or both are bottom fields, in both alternatives the previous field or frame or the next field or frame may be the current field or frame;
- the estimate for the noise level derived for a previous source field or frame.

[0014] The computation includes the following steps (cf. Fig. 2):

- a) dividing the current source field or frame into a predetermined raster of FD blocks and integrating the absolute values of the FDs over only those FD blocks which are made up exclusively of pixels classified as static in the map

of static areas;

b) translating the resulting block FDs into a first preliminary estimate of the standard deviation of the noise according to a predetermined noise model;

c) dividing the current field or frame into a predetermined raster of DFD blocks and integrating the absolute values of the DFDs over these blocks;

d) translating the resulting block DFDs into a second and a third preliminary estimate of the standard deviation of the noise according to a predetermined noise model;

e) computing a fourth preliminary estimate of the current noise level as a function of the first three preliminary estimates;

f) filtering this fourth preliminary estimate using the final noise level estimate computed for a previous field or frame to provide the final noise level estimate for the current field or frame.

A noise model assigns to a detected distribution of amounts of pixel differences in a block a corresponding noise level.

[0015] Ideally, if motion estimation is error-free, the remaining differences between the grey levels of input pixels from the two source picture blocks corresponding to, or mapped by, a motion vector must be the result of noise. The statistical distribution of the DFDs thus provides an advantageous starting point for the noise level estimation.

In practical systems, however, the accuracy of motion estimation is limited by such factors as the finite coding accuracy of the vector components, the finite spatial resolution of the source pictures, deviation of the actual scene motion from the assumed motion direction model which is usually a translative motion, and unavoidable estimation errors due to motion analysis failure, e.g. in objects containing periodic structures or in covered/uncovered areas or in pixel blocks containing a static area and smaller moving objects. The resulting motion estimation inaccuracies translate into residual DFD terms which add to the noise contribution, and thus bias the estimation of the real noise level.

[0016] Perfect motion estimation, i.e. with infinite accuracy, is nevertheless achievable on non-moving parts of the input sequence, provided that such areas exist in the current picture, which is not the case e.g. during camera panning, and that a secure method, which may or may not use motion vector information, of detecting these areas is implemented. Indeed, motion vector components in static areas are exactly zero. As a result, non-displaced frame differences or field differences between interlaced fields of identical parity, when computed on a static picture area, provide samples of pixel-wise interframe noise signal differences which are unspoiled by any residual terms stemming from motion estimation inaccuracies.

[0017] Advantageously, in the invention these two procedures are combined: one based on DFDs and the other based on FDs over static picture areas.

In situations, e.g. of camera panning, where all picture pixels are in motion and therefore no noise level estimate can be derived from the FDs, advantageously a fallback scheme can be implemented. For example, it can be decided in these cases to base the estimation solely on DFD information, or to retain the estimate computed for the previous field or frame.

[0018] In step a) the absolute values of the FD samples are integrated over predetermined pixel blocks in the current field or frame, hereafter referred to as FD blocks $FDB(i,j)$, which may or may not be overlapping. Only the FD blocks which are made up exclusively of pixels classified as non-moving in the map of static areas are used in the estimation process. For each of these FD blocks an accumulated frame difference $AFD(i,j)$ is computed as the sum of the absolute values of the FDs associated to the pixels making up the block.

[0019] It is the purpose of step b) to derive a first preliminary estimate σ_{p1} of the standard deviation of noise, expressed in grey levels, from the set of $\{AFD(i,j)\}$. This computation can be adapted to an a priori noise model. In one embodiment of the invention it is assumed that the distribution of the absolute values of the FDs associated with static pixels is such that its mean $m_{|FD|}$ is proportional to the standard deviation σ of the noise level to be estimated:

$$m_{|FD|} = k \cdot \sigma.$$

[0020] This assumption holds in particular when the samples of the source noise are spatially and temporally uncorrelated and follow a Gaussian distribution, in which case k is found to equal $2/\sqrt{\pi} \approx 1.13$. In one embodiment of the invention, k is set to this value. The mathematical expectancy of the AFDs, which can be approximated by the arithmetic average $\{AFD(i,j)\}$ of the $AFD(i,j)$ over the static blocks within the current field or frame, is given by $N_{FDB} \cdot m_{|FD|}$, where N_{FDB} represents the number of pixels in an FD block.

[0021] A good approximation to σ can therefore be derived as:

$$\sigma_{p1} = \{AFD(i,j)\} / (k \cdot N_{FDB})$$

In step c), which is similar to step a), the absolute values of the DFD samples are integrated over predetermined blocks in the current field or frame, hereafter referred to as DFD blocks $DFDB(i,j)$. These blocks may or may not be overlapping.

For each DFD block DFDB(i,j) an accumulated DFD referred to as {ADFD(i,j)} is computed as the sum of the absolute values of the DFDs associated to the pixels making up the block.

In step d), which is similar to step b), the set of {ADFD(i,j)} is translated into a second σ_{p2} and a third σ_{p3} preliminary estimate of the standard noise deviation expressed in grey levels. The derivation of σ_{p2} is identical to that of σ_{p1} with the exception that the set of {AFD(i,j)} is replaced by the set of {ADFD(i,j)}. Let N_{DFDB} be the number of pixels in a DFD block and $\{ADFD(i,j)\}$ be the average of the ADFDs for the current field or frame. Then σ_{p2} is computed as:

$$\sigma_{p2} = \{ADFD(i,j)\} / (k * N_{DFDB})$$

[0022] Unlike for FDs, however, the estimation of noise level based on DFDs may be biased by residual terms resulting from motion estimation imperfections as explained above. This is likely to occur if the processed fields or frames contain areas with high spatial gradients. In order to improve the robustness of the proposed method, a third preliminary estimate σ_{p3} is derived from the minimum rather than the average of the ADFD(i,j):

$$\sigma_{p3} = \min_{i,j} (ADFD(i,j)) / (k * N_{DFDB})$$

[0023] In step e) a single preliminary estimate σ_p is derived from σ_{p1} , σ_{p2} and σ_{p3} . First, the ratio $r = \sigma_{p2}/\sigma_{p3}$ is thresholded to determine which preliminary estimates should be used.

A value of r above a predetermined threshold T_r set to a value in the range between "1" and "5", preferably to the value "2" in one embodiment of the invention, indicates a large variety of textures and therefore a significant proportion of high gradient areas in the source picture. In that case σ_{p2} is deemed to be unreliable and the preliminary estimate σ_p is computed from σ_{p1} and σ_{p3} only.

Conversely, if r falls to or below T_r , indicating consistency of the estimates computed from the block DFDs, σ_{p2} and σ_{p3} as well as σ_{p1} are used.

Advantageously σ_p is derived as:

$$\begin{aligned} \sigma_p &= (\sigma_{p1} + \sigma_{p3})/2 & \text{if } \sigma_{p2}/\sigma_{p3} > T_r \\ \sigma_p &= \text{median}(\sigma_{p1}, (\sigma_{p1} + \sigma_{p2})/2, \sigma_{p3}) & \text{if } \sigma_{p2}/\sigma_{p3} \leq T_r \end{aligned}$$

where median() means a 3-tap median filter.

[0024] Since fast variations of the actual noise level in a broadcast image sequence are very unlikely, in step f) a temporal low-pass filter is applied to σ_p to further improve the robustness of the noise level estimation. The final estimate σ of the standard deviation of the noise level is computed from σ_p and from the noise level estimate σ_{prev} of a previous frame or field or field of corresponding parity as:

$$\sigma = \text{median}(\sigma_{prev} - \Delta v_{low}, \sigma_p, \sigma_{prev} + \Delta v_{high})$$

[0025] Δv_{low} and Δv_{high} are predetermined constants that specify the maximum variations of the estimated noise level variance from one estimation cycle (e.g. field or frame) to the next. In one embodiment of the invention Δv_{low} and Δv_{high} are set to about "1" and about "0.25" grey levels, respectively.

[0026] The threshold values given in this application are based on an 8-bit representation of the pixel values. If these pixel values have a different resolution, the threshold values should be adapted accordingly.

[0027] It may happen that motion estimation is performed on couples of fields or frames that are not consecutive, in which case the current source picture for which the noise level is estimated may differ from the pictures used for motion estimation. This is the case e.g. in an MPEG2 encoding scheme if the current frame is a B-frame.

[0028] One or both of said fields or frames used for determining the pixel value differences FD concerning static picture areas can be different from one or both of said fields or frames used for determining the pixel value differences DFD concerning couples of motion-compensated blocks or couples of blocks mapped by their associated motion vector.

[0029] One may use all blocks of the active part of the fields or frames concerned for the noise level computation. However, it is also possible to not consider pixel blocks which are located at the borders of the active picture part, in

particular because the motion information for such blocks may be less reliable. It is also possible to limit further the number of blocks considered per picture.

Claims

- 5 1. Method for estimating the noise level (σ_{p2} , σ_{p3} , σ) for a current source field or frame (IF) of a video sequence, based on the differences (DFD, FD) between pixel values of blocks in a previous field or frame (PF) and corresponding pixel values of corresponding blocks in a future field or frame (NF), wherein either said previous (PF) or said future (NF) field or frame can be said current field or frame (IF) itself, characterised in that at least one block of each corresponding couple of blocks
10 is a motion-compensated pixel block or is mapped to the other block by an associated motion vector estimate.
2. Method according to claim 1, wherein in addition static picture areas are determined and the differences (FD) between pixel values of blocks in a static picture area of a previous field or frame (PF) and corresponding pixel values of corresponding blocks in a future field or frame (NF) are used to estimate a further noise level estimate (σ_{p1}) which is combined with said noise level estimate (σ_{p2} , σ_{p3}) in order to form a final noise level estimate (σ), wherein said previous (PF) and/or said future (NF) field or frame used for the evaluation of said differences (FD) between pixel values of a block in a static picture area can be different from said previous (PF) and/or said future (NF) field or frame used for the evaluation of the differences (DFD) concerning said motion-compensated pixel blocks or said mapped blocks.
15 20
3. Method according to claim 1 or 2, wherein the amount values of said differences (DFD, FD) between pixel values become accumulated (ADFD, AFD) for each block.
25
4. Method according to any of claims 1 to 3, wherein said blocks are overlapping.
5. Method according to any of claims 1 to 4, wherein for said noise level estimate two estimates (σ_{p2} , σ_{p3}) are calculated, wherein the first one (σ_{p2}) is derived from the average ($\{ADFD(i,j)\}$) of the accumulated block pixel difference values (ADFD) for the current field or frame and wherein the second one (σ_{p3}) is derived from the minimum ($\min(ADFD(i,j))$) of the accumulated block pixel difference values for the current field or frame.
30
6. Method according to claim 5, wherein the final noise level estimate (σ) is the mean of said further noise level estimate (σ_{p1}) and the second one (σ_{p3}) of said noise level estimates (σ_{p2} , σ_{p3}), if the ratio between the first one (σ_{p2}) and the second one (σ_{p3}) of said noise level estimates is greater than a predetermined threshold (T_r), in particular about "2",
35 and wherein the final noise level estimate (σ) is the median of said further noise level estimate (σ_{p1}) and of the mean of this further noise level estimate (σ_{p1}) and the first one (σ_{p2}) of said noise level estimates and of said second one (σ_{p3}) of said noise level estimates, if the ratio between the first one (σ_{p2}) and the second one (σ_{p3}) of said noise level estimates is equal to or smaller than said predetermined threshold (T_r).
40
7. Method according to claim 6, wherein said final noise level estimate (σ_p) becomes median filtered together with a noise level estimate for a previous frame or field (σ_{prev}) from which a first predetermined constant (Δv_{low}) is subtracted and with said noise level estimate for a previous frame or field (σ_{prev}) to which a second predetermined constant (Δv_{high}) is added, in order to form a final output noise level estimate (σ).
45
8. Method according to claim 7, wherein said first and second predetermined constants (Δv_{low} , Δv_{high}) specify the maximum variations of the estimated noise level variance from one estimation cycle to the next.
- 50 9. Method according to claim 7 or 9, wherein said first predetermined constant (Δv_{low}) has a value of about "1".
10. Method according to any of claims 7 to 9, wherein said second predetermined constant (Δv_{high}) has a value of about "0.25".
- 55 11. Method according to any of claims 1 to 10, wherein in situations where all or nearly all pixels of a picture are in motion, a fallback noise level estimation is carried out and that estimation is based solely on the determination of pixel value differences (DFD) concerning motion-compensated interpolated pixel blocks or blocks mapped by an associated motion vector estimate, or

EP 0 957 367 A1

is based on a noise level estimate computed for a previous field or frame.

5

10

15

20

25

30

35

40

45

50

55

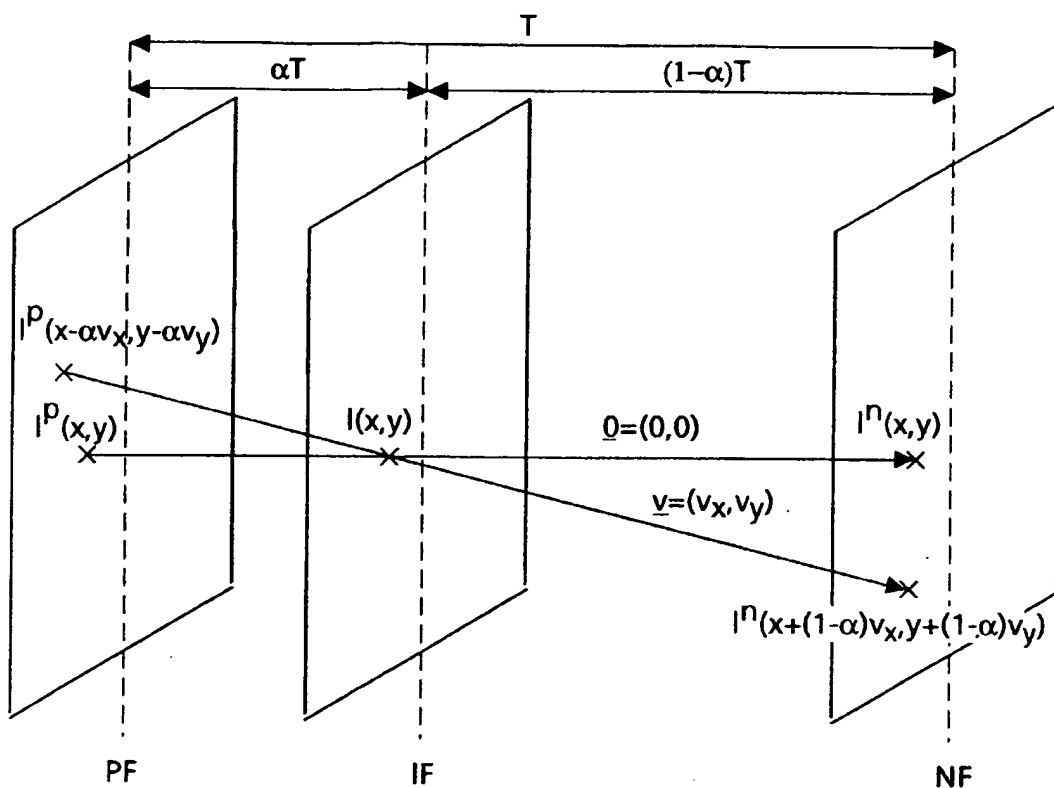


Fig.1

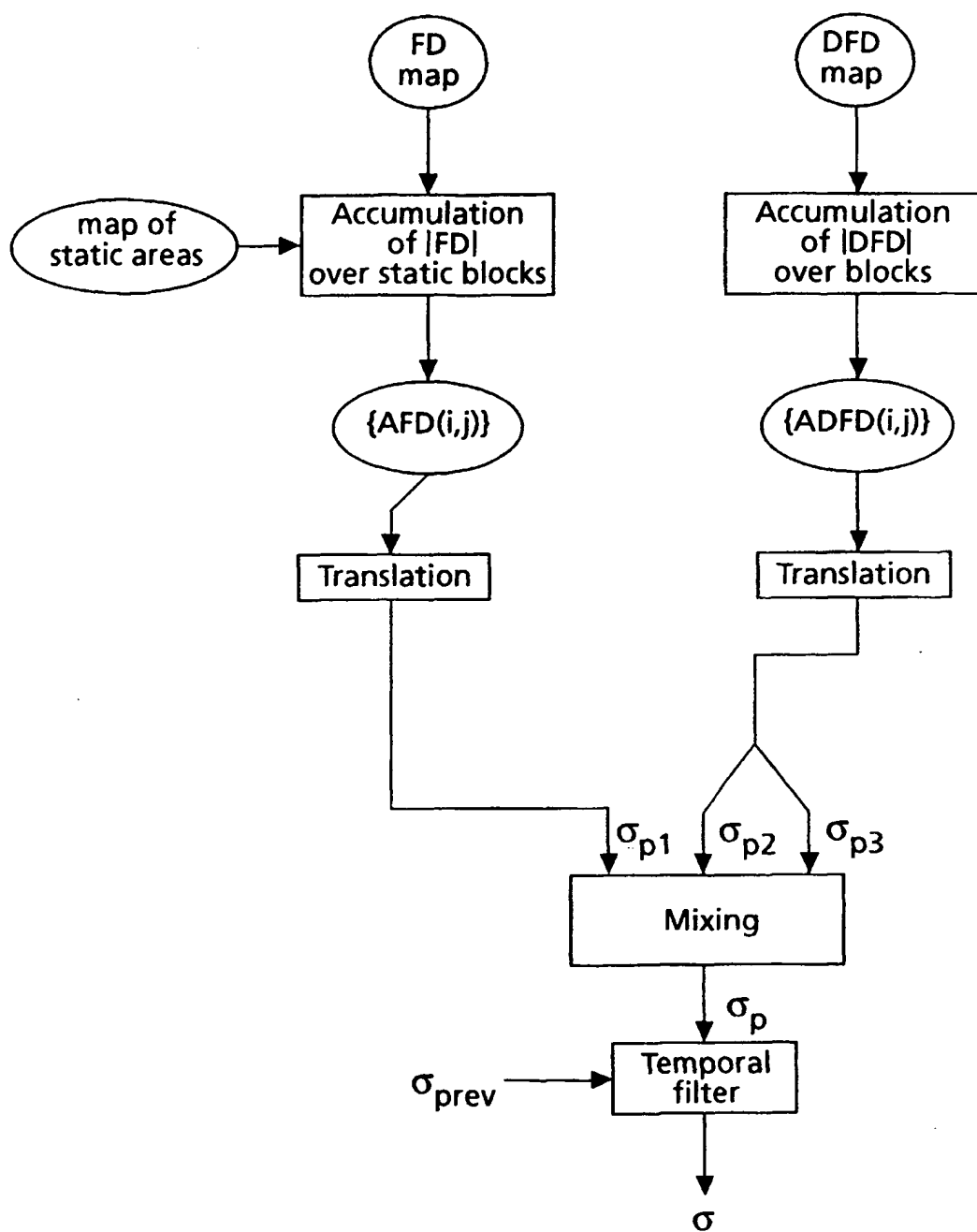


Fig.2



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 98 40 0917

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
A	EP 0 735 747 A (THOMSON MULTIMEDIA SA) 2 October 1996 * page 6, line 55 - page 7, line 9 *	1	G01R29/00 H04N17/00
A,D	EP 0 562 407 A (THOMSON CONSUMER ELECTRONICS) 29 September 1993		
			TECHNICAL FIELDS SEARCHED (Int.Cl.6)
			H04N
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 11 September 1998	Examiner Yvonnet, J
<p>CATEGORY OF CITED DOCUMENTS</p> <p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p> <p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document</p>			

EPO FORM 1503 03 82 (P04C01)